

Pilot Communication and Execution Latencies 1

RUNNING HEAD: Impact of UAS Pilot Communication and Execution Latencies

Impact of UAS Pilot Communication and Execution Latencies on Air Traffic Controller's

Acceptance of UAS Operations

Kim-Phuong L. Vu, Dan Chiappe, Gregory Morales, Thomas Z. Strybel

Center for Human Factors in Advanced Aeronautics Technologies and California State

University Long Beach, Long Beach, CA

Vernol Battiste

San Jose State University at NASA Ames Research Center, Moffett Field, CA

Jay Shively

NASA Ames Research Center, Moffett Field, CA

Timothy J. Buker

SAIC, Washington, D.C.

CORRESPONDENCE TO:

Kim-Phuong L. Vu

Department of Psychology, California State University Long Beach

1250 N Bellflower Blvd., Long Beach, CA, 90840

E-MAIL: kim.vu@csulb.edu

Abstract

The integration of Unmanned Aircraft Systems (UAS) into the National Airspace System (NAS) requires that UAS meet or exceed the safety requirements established for conventional aircraft, and for the UAS pilots to interact with air traffic controllers (ATCos) in an acceptable manner. UAS have several characteristics that differentiate them from conventional aircraft, including the possibility of greater latencies associated with remote pilot communication and command execution. The goal of the present study was to determine how adding delays to UAS pilot communications and command executions affect ATCos' interactions with UAS and conventional aircraft. Six previously certified radar controllers and two currently certified radar controllers were recruited as participants to manage traffic in a simulated sector with conventional traffic and one UAS flying in it. The UAS pilot verbal communication and execution latencies were varied in separate scenarios to include an additional delay that was either short (1.5 s) or long (5 s), and constant or variable within each scenario. We measured both UAS and conventional pilots' verbal communication and execution initiation latencies, and obtained ATCos' acceptability ratings for the different delay conditions. Also examined were the number of communication step-ons created by the additional communication delays implemented in the UAS control station, as well as other measures of the ATCo-pilot interactions. We found ATCos rated UAS pilot verbal communication latencies to be acceptable when the latencies were short rather than long and that acceptability ratings often reflect broader features of the sectors being managed. Implications of these findings for UAS integration in the NAS and limitations of the present study are discussed.

MAIN TEXT WORD COUNT: 9,860 words

3 FIGURES

Impact of UAS Pilot Communication and Execution Latencies on Air Traffic Controller's Acceptance of UAS Operations

I. Introduction

1.1 Statement of the problem

In 2012, the U.S. Congress passed the FAA Modernization and Reform Act. It calls for a plan to integrate Unmanned Aircraft Systems (UAS) into the National Airspace System (NAS). For UAS to be allowed to operate with unrestricted access, they will be required to “act and respond as manned aircraft do” (ICAO, 2011, p. 5). Research is required to determine what standards and regulations should be applied, modified, or newly developed (Ambrosia, Cobleigh, Jennison, & Wegener, 2007; Askelson et al., 2013; Blickensderfer et al., 2012; Dillingham, 2013). FAA regulations, for example, require that pilots need to comply promptly with air traffic controller (ATCo) commands (14CFR 91.123). However, the regulations do not quantify what an acceptable response latency is, stating only that they must not compromise the safe separation of aircraft. To begin addressing this issue, the present study examined how UAS pilot latencies in verbal communication and command execution affect sector-wide ATCo performance, communication, and controller acceptability ratings of both conventional aircraft and UAS operating in a NAS environment.

A key challenge facing the integration of UAS stems from the fact that these aircraft differ from conventional ones in many important respects. For example, UAS pilots face the unique challenge that their aircraft can fall prey to lost-link situations, where the pilots cannot directly control their aircraft. If this happens in an integrated airspace, it can significantly add to the workload of ATCos (Kamienski, Simons, Bell, & Estes, 2010). Other differences between these types of aircraft also affect the time UAS require to respond to ATCo commands (Apaza &

Kubat, 2012; Askelson, et al., 2013). For example, UAS tend to be smaller, slower, and have very different control interfaces. Most crucially, UAS pilots are not co-located with their aircraft (Williams, 2007). The pilots therefore lack many of the sensory cues (e.g., ambient vision, vestibular, and acoustic information) available to conventional pilots to make fast assessments about the state of their aircraft and to, for example, see and avoid other aircraft (14CFR 91.113, as currently interpreted for UAS). UAS pilots also have difficulties in scanning the environment around their aircraft because the cameras, if available, typically have limited spatial resolution and a small field of view (Merlin, 2013). At least at present, then, UAS have poor “sense and avoid” capabilities, leading UAS pilots to take longer to determine whether they can safely carry out commands issued by ATCos (Dillingham, 2013; McCarley and Wickens, 2005; Vu & Chiappe, in press; Williams, 2012).

Verbal communication can also take longer for UAS depending on the communication infrastructure employed. Satellite communication links, as opposed to the UHF/VHF radio communications used by pilots of conventional aircraft, lengthen the time to verbally respond by increasing audio set up and signal propagation times (Dillingham, 2013; Wourms, Ogden, & Metzler, 2001). Indeed, long latencies associated with voice based communications that make use of satellite technology can increase ATCo workload, and negatively affect their performance by increasing the number of stepped on communications (Nadler, Mengert, DiSario, Grossberg, and Spanier, 1993; Sollenberger, McAnulty, and Kerns, 2003). Satellite communication delays are also a concern for the proposed shift towards Data Comm, envisioned as part of the NextGen air traffic management system. Kerns (1991), in a review of Data Comm communications issues, observed that the total transaction time for Data Comm is on average twice as long as that for voice communications, although communication via Data Comm is more precise and concise.

Moreover, Data Comm may alter the sequence of communications between pilots and controllers: With Data Comm, pilots have more flexibility in when they respond to ATCo instructions compared to voice, because the latter requires a prompt response as spoken clearances can be quickly forgotten (Helleberg and Wickens, 2003).

Pilot execution latencies can also be greater for UAS than conventional aircraft depending on the type of control interface involved (Askelson et al., 2013). For example, some stick, throttle and rudder control inputs, such as those in the Predator-B, may allow for faster response execution than some systems that involve supervisory control (i.e., human-on-the-loop), such as the Global Hawk (Williams, 2007). In supervisory control systems, pilots issue high-level commands to aircraft, which then perform them autonomously (Sheridan, 1992). Although the interface design of the ground control station could be made to allow very quick responses, current interfaces often require pilots to enter commands by engaging in time-consuming operations, such as clicking through computer screens and pull-down menus, and making keyboard entries to carry out even simple maneuvers such as changing altitudes, airspeeds, or waypoint destinations (Merlin, 2013).

In short, UAS have various characteristics that can lengthen the time they require to respond to commands issued by ATCo in comparison to conventional aircraft. However, this does not *ipso facto* mean that UAS response latencies will prevent safe operations in the NAS. After all, there is a great deal of variability with which conventional aircraft pilots respond to ATCo instructions (Cardosi, 1993). What it does mean, however, is that empirical research is required to determine what UAS-generated delays will affect air traffic management performance, workload and acceptability as these are considered the most sensitive indicators of delay effects in the NAS (e.g., Sollenberger et al., 2003).

1.2 Purpose of the study

The end-to-end response time of pilots to ATCo verbal commands is known as the “measured response” (MR; Shively, Vu, & Buker, 2013; Vu et al., 2013; Ziccardi et al., 2013). It can be broken down into components (see *Figure 1*) that include the following: (1) *Pilot verbal latency (MR1)*; the lag between the end of an ATCo’s verbal clearance and the beginning of the pilot’s read back. (2) *Pilot execution latency (MR2)*; the lag between the end of the ATCo’s clearance and when a pilot starts to execute the maneuver. (3) *Aircraft response latency (MR3)*; the lag between the pilot entering a command in the control interface and the aircraft acting in response to that command. (4) *Display visibility latency (MR4)*; the time for the maneuver to be available on the ATCo radar screen after the aircraft has started its maneuver. In the present study, we manipulated the UAS pilot verbal latency and execution latency components (MR1 and MR2) to examine their effect on ATCo sector performance, communication, and acceptability ratings. The purpose was therefore to determine the impact of delays in verbal responding and initiation of command executions on ATCo performance and their acceptability ratings of pilot responses.

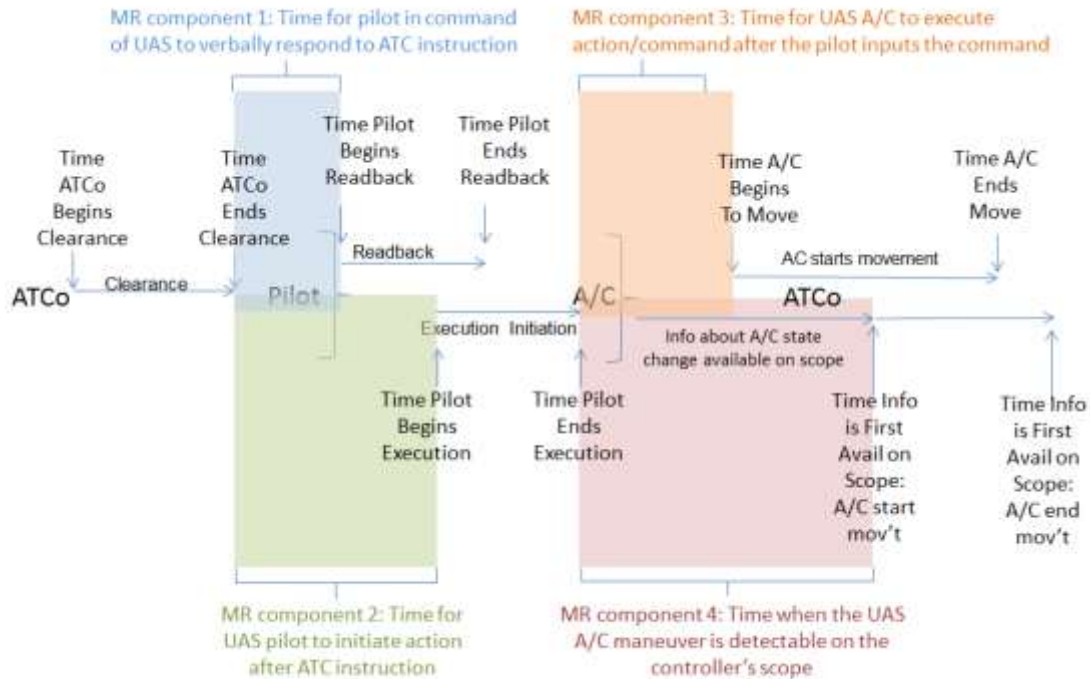


Figure 1. Illustration of the measured response (MR) components in the course of the ATCo-Pilot communication interaction.

1.3 Literature review

Askelson et al. (2013) conducted measured response research using UAS. They used a Predator (MQ-9) simulator flying in stick and rudder mode to execute commands issued by ATCos. The commands (heading, altitude, and speed changes) were issued during different phases of flight. Askelson et al. (2013) measured what they referred to as the “response time” of the UAS. This is the lag between the end of the ATCo’s clearance and when it becomes apparent on the ATCo display that the UAS is complying with the command. They also measured the maneuver completion time—the time between the end of the ATCo clearance and when the desired result was achieved by the UAS. They found average response times of 11.85 s across all maneuvers, and completion times that averaged 71.20 s (ranging from 58.70 s to 88.25 s depending on the clearance). Although these findings do provide important information

regarding overall response times of one type of UAS, they are limited with respect to the purpose of the present study. Specifically, Askelson et al. (2013) did not examine the latency in verbal responding by UAS pilots (MR1) or when the UAS pilots began executing maneuvers (MR2). Moreover, they did not examine the ATCos' acceptability ratings of the latencies that they obtained.

Likewise, research on conventional aircraft, while examining some of the measured response components, also has important limitations with respect to the purpose of the present investigation. For example, Cardosi (1993) examined recordings of verbal interactions between pilots and controllers from three en route sectors. She found mean pilot verbal latencies ranging from 2.67 s ($SD = 6.25$ s) to 3.31 s ($SD = 4.80$ s), with total transaction times of approximately 10 s. More recently, Smith (2008) conducted a similar study, using recordings from TRACON sectors. Given that these are generally busier sectors it is not surprising that she found that the verbal latencies between pilot and controller responses were shorter, with an average of about 1 s. Nonetheless, these findings cannot be used to identify what is an acceptable pilot verbal latency. This is because these studies did not have ATCos rate acceptability of the latencies in verbal responding by pilots.

The existing literature offers even less guidance in the case of acceptable latencies in command execution, where very little research has been conducted. A notable exception is a study carried out by Consiglio, Hoadley, Wing, Baxley and Allen (2008). Their simulation study manipulated delays in pilots executing commands issued by an Airborne Separation Assistance System (ASAS) flight deck tool. Using losses of separation (LOS) as their dependent variable, they found that the ASAS system was able to maintain high levels of safety, except when pilot execution latencies were greater than 90 s and traffic densities were at extreme levels. Although

the results from this simulation are useful for guiding future research, the values identified do not allow us to determine acceptable pilot execution latencies. This is because the simulated pilots were not responding to ATCo communications, but rather to a flight deck automation tool.

1.4 Research questions and hypotheses

In the current study we manipulated whether UAS pilot verbal and execution initiation responses were “short” or “long.” This was done by inserting a delay of 1.5 s or 5 s to their verbal responses and by inserting a delay of 1.5 s or 5 s prior to the beginning of their command executions. These values were chosen based on the acceptability ratings collected by Shively et al. (2013), in which controllers issued clearances serially to UAS pilots and then rated the acceptability of their latencies in responding. In addition to manipulating pilot verbal and execution delays, the present study also varied whether the delays were constant or variable within a scenario. In the latter, some of the delays in verbal responding (and command execution) were short, some long within the same scenario. In the former, the delays were held constant throughout a scenario. In the context of UAS operation in the NAS, the predictability of the response latencies is important to investigate because UAS may have different communication latencies within the same sector, depending on whether pilots are communicating with ATCos via VHF/UHF radio or through a satellite communication link. Execution latencies may also differ depending on whether the UAS pilots are controlling their aircraft via line of sight or satellite links, the latter yielding longer latencies (Dillinger, 2013; Merlin, 2009).

We carried out this study to address the following research questions and to test their accompanying hypotheses:

- (1) What is the effect of delays in verbal responding by UAS pilots on the ATCo acceptability ratings of the resulting MR1 components of all aircraft in the sector?

Hypothesis 1: In general, ATCos should rate the MR1 to be more acceptable in the short delay condition than in the long delay condition, and the short delay conditions should generally be acceptable to ATCos. This is because in the short delay condition, the MR1 component is likely to be close to the MR1 values obtained by Cardosi (1993) for conventional aircraft.

Hypothesis 2: Because the delays are added only to the UAS pilot communications, the acceptability ratings for the conventional pilots should be higher than those of the UAS pilots.

(2) What is the effect of the predictability of the UAS-pilot verbal latencies (MR1) on ATCo acceptability ratings of MR1?

Hypothesis 3: ATCo acceptability ratings of MR1 will be higher when delays are constant than variable. Although much research has shown that adding signal delays can negatively affect performance (e.g., Clark, 2003; Ferrell, 1965; Held, Efstathiou, & Greene, 1972; Sheridan, 1992; 1993), we maintain that the negative effects of delays will be more pronounced when they are unpredictable to operators. This is because when a lag is constant, operators may be able to develop a strategy for managing it; an ATCo, for example, may learn to wait for a specific amount of time before repeating a command, but this strategy could not be employed if the delay is unpredictable. Delay unpredictability has been found to be important to coordination efforts, as the uncertainty of aircraft arrival times is a major source of operational costs for the airline industry (Agbolosu, Millner, Baden, Coville, & Mondolini, 2012).

(3) What is the effect of delays in initiating command executions on the ATCo acceptability ratings of the resulting MR2 components of all aircraft in the sector?

Hypothesis 4: ATCo acceptability ratings should be higher in the short delay condition than the long delay condition.

Hypothesis 5: Differences in acceptability of MR2 between the long and short conditions should be smaller than for the delays in verbal responding (MR1). That is, delays in verbal responding should be more salient for ATCos than delays in command execution. This is because ATCos typically wait to get verbal confirmation from a pilot after they issue a clearance, but often continue to work with other aircraft in the sector before checking that the command is being executed correctly.

Hypothesis 6: The acceptability ratings for the MR2 components will be higher for conventional aircraft than UAS because the delays were added only to the latter.

(4) What effect does the predictability of the UAS pilot execution initiation latencies (MR2) have on the ATCo acceptability ratings of MR2?

Hypothesis 7: Constant delays in response initiation by UAS pilots will receive higher acceptability ratings from ATCos than variable delays. However, because verbal delays by pilots are likely to be more salient than latencies in command execution, it is likely that the differences between constant and variable MR2 latencies will be smaller than the differences for MR1.

(5) What is the effect of verbal and execution response latencies on ATCo performance?

Hypothesis 8: We expect ATCo performance to remain high regardless of our manipulations given that the participants are highly skilled. However, ATCo safety performance, efficiency, and communication will be better in the short UAS verbal delay condition than in the long verbal delay condition. ATCo performance will also be better under short than long execution initiation delay conditions.

Hypothesis 9: ATCo performance will be better in the constant delay conditions than in the variable delay conditions. This is because they will be able to develop strategies for handling the delays when they are constant within a scenario, but will be unable to do so when they vary.

(6) What, if any, other factors influence ATCo ratings of the acceptability of delays in verbal responding and command execution?

The ATCos were asked to rate the latencies in verbal responding and command execution for their acceptability for both UAS pilots and conventional aircraft pilots after each scenario. It is possible that these ratings were based on more than the delays themselves, and if so, what factors influence acceptability ratings? Acceptability ratings could be based on various ATCo performance factors, such as the number of step-ons, losses of separation (LOS), and measures of efficiency, because of the contextual nature of air traffic management. For example, verbal latencies are shorter in TRACON than in en route sectors. This is because the airspace is much more complex in the former, and actions need to be carried out more quickly. What is acceptable in one case is not likely to be acceptable in another. Likewise, a delay that is acceptable at night when there are fewer aircraft in the sector may not be appropriate during peak operations during the day. Therefore, we examined the relationships between acceptability ratings and other controller performance measures.

II. Method

Participant and Confederate Demographics

Eight air traffic controllers (ATCos) participated in the simulation. Six of the participants were retired (≤ 7 years since retirement) but previously radar-certified, and two were current, off-

duty volunteers. The controllers all had experience in air traffic management in civilian facilities, and four also reported experience in military air traffic management. Altogether, they reported an average of 28 years ($SD = 4.8$ years) of air traffic management experience in military or civilian facilities. The controllers all had prior experience with Los Angeles (ZLA) airspace that was used in this simulation. In addition, all of the controllers had participated in prior, unrelated studies using our simulation facility. However, none had participated in simulations involving UAS flying in an integrated airspace. The present study was scheduled over two consecutive days, and participants were compensated \$960 for their time. All participants completed the study, but the data from one participant was excluded for non-compliance with some of the experimental procedures.

All aircraft (UAS and conventional) were flown by experimental confederates. Although they were highly trained on the piloting software and procedures used in the study, none were IFR certified pilots. As such, they are referred to as “pseudopilots” throughout the paper.

Design

Independent variables. The present study had three independent variables. These were, (1) the UAS-pseudopilot verbal delay (Short or Long), (2) the UAS-pseudopilot execution initiation delay (Short or Long), and (3) the predictability of the delays (Constant or Variable). For both Communication Delay and Execution Delay conditions, a delay of either 1.5 s (Short Delay) or 5 s (Long Delay) was inserted before the UAS-pseudopilot communications with ATCos or prior to the UAS pilot initiating the execution of a clearance. For the constant conditions, combining the communication and execution delays produced four conditions: Voice Delay-Short/Execution Delay-Short, Voice Delay-Short/Execution Delay-Long, Voice Delay-Long/Execution Delay-Short, and Voice Delay-Long/Execution Delay-Long, tested in separate

scenarios. In terms of the Predictability variable, we compared findings from the Constant Delay scenarios that employed one of the four delay combinations throughout the scenario with the variable delay trials that employed two repetitions of each delay combination within each scenario.

The study did not examine the aircraft response latencies (MR3) because our simulation software does not allow adjustments to these parameters. In our simulations, aircraft respond immediately once pilots enter commands. We measured the MR4 component, display visibility latency, but did not manipulate it as it was relatively constant throughout the simulation at approximately 5 s, and was influenced by the update rate set by the DSR display and the recording software.

Dependent variables. In the present study, ATCos were not rating the acceptability of the imposed delays themselves, but rather the resulting MR1 and MR2 components that included the delays in addition to the pilots' response time. As a result, we first had to measure the following components: Verbal Latency (MR1), the time between the end of an ATCo's verbal clearance and the beginning of the pilot's read back, and Execution Initiation Latency, (MR2), the time between the end of the ATCo's clearance and the pilot beginning to execute the clearance. We calculated these for both UAS pilots and conventional aircraft pilots, though the delays were only added to the former. For conventional aircraft, MR1 and MR2 were measured using a sample of communications for each scenario and participant ATCo. The communications were equal in number to those of UAS pilots and ATCos, and were taken from similar points in the scenarios to likely reflect equivalent ATCo workload.

In addition, to answer the research questions outlined above, and to test their attendant hypotheses, the following dependent variables were measured:

(1) ATCo acceptability ratings. Following each scenario, participants were presented with the following questions: “How acceptable were the delays in verbal responding by the UAS pilots?”, “How acceptable were the delays in verbal responding by conventional aircraft pilots?”, “How acceptable were the delays in command execution by the UAS pilots?”, and “How acceptable were the delays in command execution by the conventional aircraft pilots?” Each question was presented with a scale ranging from 1 (very unacceptable) to 7 (very acceptable). These subjective measures of acceptability are similar to the rating scales used by Rantanen, McCarley and Xu (2004) and by Sollenberger, McAnulty and Kearns (2003). The former, for example, found that acceptability ratings were sensitive to variations in audio set up delays and pilot verbal delays in a part-task simulation.

(2) Losses of separation (LOS). The average number of LOS per condition was used as a measure of the safety with which ATCos managed traffic in their sector. An LOS was defined as two or more aircraft coming within 5 nm laterally and within 1000 ft vertically of each other. Using LOS as a measure of safety is justified by the fact that the main goal of air traffic controllers is to maintain safe separation of aircraft in their sector. Moreover, it is commonly used for this purpose (see, e.g., Prevot, Homola, Martin, Mercer, & Cabrall, 2012; Vu et al., 2012).

(3) Distance aircraft travel through the sector. This was used as a measure of efficiency, calculated by the average number of nautical miles (nm) travelled through the sector by each of the conventional aircraft. This is a common measure of flight path efficiency, as one of the goals of air traffic management is to get aircraft through the sector using as efficient a route as possible (e.g., Prevot et al., 2012; Vu et al., 2012). This is important for the conservation of fuel and to allow aircraft to get to their destination on time.

(4) Stepped on communications (step-ons). These were defined as any time two or more communications overlapped with each other, making them unintelligible. This included communications between ATCos and both conventional aircraft pilots and UAS pilots. These were used as a measure of sector performance because of their link with serious aviation accidents and incidents, including the 1977 Tenerife crash involving two 747s. Other researchers have used step-ons as a measure of sector safety including Nadler et al. (1993), who found that a greater number of step-ons resulted from longer audio set up latencies, particularly when traffic density was high (see also Sollenberger et al., 2003).

Apparatus

To ensure that communications between operators were only made via radio, the simulation was conducted in three rooms. Each room contained workstations for one of the three operator roles— participant ATCos, UAS pseudopilots and conventional-aircraft pseudopilots. For ATCos and conventional-aircraft pseudopilots, the simulation used the Multi Aircraft Control System (MACS; Prevot, 2002), a medium-fidelity simulation architecture. MACS simulated a DSR display of sector ZLA-20 for ATCos, and a pseudopilot flight deck display for conventional-aircraft pseudopilots. The Multiple UAS Simulator (MUSIM; for a detailed description see Fern and Shively, 2009) Ground Control Station was used to fly the UAS. The UAS pseudopilots controlled their aircraft by altering the altitude and waypoints (point-to-point navigation) in MUSIM. The UAS callsign was “PD-1” in all scenarios. Confederate UAS and conventional-aircraft pseudopilots were students in CHAAT with extensive training on each role.

Two parallel worlds were run simultaneously. The controllers were seated in front of a radar display and had a ‘Mission Control’ display off to the side, which was used to instruct the controller on the desired waypoint to send the UAS. The ATCo, conventional-aircraft

pseudopilots, and UAS pseudopilots communicated with each other through a VoiceIP system using push-to-talk headsets. The voice client for the UAS station was modified in two ways. First, a mechanism for inserting a delay in the UAS-pilot transmissions was developed. Second, the voice system was modified to produce blocked transmissions if any operator stepped-on the transmission of another operator. In this case, both transmissions became unintelligible to everyone listening on the frequency until only one operator was speaking. The voice system also logged the number and duration of step-ons.

To implement the delays in execution initiation, the UAS pseudopilots activated a countdown timer of 1.5 s or 5 s after acknowledging the clearance. For the variable condition, the timer varied between short and long randomly within the required number of delays. The UAS pseudopilots were given extensive practice on these procedures prior to the study, and no problems were reported by pseudopilots or their experimenter observers in carrying them out during the simulation. During the variable-delay conditions, the voice and execution delays were counterbalanced within the scenario and each delay combination was presented twice.

Procedure

The simulation was run over two consecutive days. After giving informed consent, participants completed demographic questionnaires and were briefed on the operational environment and specific procedures for the simulation. Following the briefing, the ATCos engaged in a training session that consisted of four, 30-minute practice scenarios. The first three scenarios did not include the UAS, and were designed to re-familiarize the ATCos with the airspace and traffic flows used in the simulation. The UAS was introduced in the final practice trial, but it did not produce any delays in communication or execution. A lengthy training session was deemed necessary for the following reasons: First, although all ATCos had

experience working ZLA-20, those that were retired had not been actively managing traffic in that sector. Second, the ATCos also needed practice working the sector with a UAS present, as the latter aircraft flew a route that could interfere with normal flows of traffic in the sector. After the training session, participants were allowed to ask questions about any of the simulation procedures.

The experimental scenarios were run in the afternoon of the first day and morning of the second day. Each experimental scenario was 40 minutes in duration, with about 50 AC entering the sector during the entire scenario and at least 8 AC in the sector at any given time. During the experimental trials, the ATCos were instructed to give priority to arrivals into LAX, and to ensure that arrival aircraft left the sector at an altitude of 10,000 ft and a speed of 250 kts. The controllers were told that a Letter of Agreement with the Center was in effect, requesting accommodation of requests from either Mission Control or the UAS ground station. The UAS flew in the sector for the entire scenario in a triangular pattern at an average speed of 120 kts, and altitudes between 10,000 and 16,000 ft. These flight plans ensured that the UAS crossed the arrival streams several times during a scenario.

ATCos received audio alerts and text messages on the Mission Control station indicating a new waypoint for the UAS (e.g., "PD-1 proceed direct EDITS"). The ATCo then issued the clearance when appropriate. UAS pseudopilots requested altitude changes (e.g., "LA Center, PD-1, request descent 1- 4 thousand"). Eight planned requests for PD-1 were completed in each scenario, four initiated by the controller and four by the UAS pseudopilot. ATCo- and UAS-initiated communications occurred every 4-5 min in alternating orders, beginning 2-3 minutes into the scenario. Note that the UAS station produced only transmission delays, no receiving delays. The ATCo and conventional-aircraft-pseudopilot stations produced neither type of delay.

Because we were interested in capturing the MR components, we did not include additional secondary tasks, such as online measures of acceptability, workload, or situation awareness, due to their potential intrusiveness (see, e.g., Pierce, 2012). At the end of each trial, controllers rated the acceptability of the verbal and execution response latencies (MR1 and MR2) for both UAS and conventional aircraft. In addition, situation awareness (SART; Taylor, 1990) and workload (NASA-TLX; Hart & Staveland, 1987) measures were obtained. The SART and NASA-TLX metrics did not yield any significant effects and will not be discussed further. After finishing all 8 experimental trials, the ATCos completed a post-simulation questionnaire and participated in a debriefing session.

III. Results and Discussion

3.1 Measured response components

In presenting our findings we first describe the effect of the inserted delays on MR1 and MR2 of UAS and conventional aircraft (see Vu et al., 2013, for a detailed analysis of the MR1 and MR2 components as a function of the delay type and predictability). Not surprisingly, the UAS pilot verbal and execution initiation response times are longer than those of conventional aircraft pilots, $Fs(1,12) > 23.4$, $ps < .001$. The means and acceptability ratings for these components are listed in *Table 1*.

Table 1. Mean Pilot Verbal and Execution Delays (in secs; standard deviations in parentheses) and Mean ATCo Acceptability ratings (1 = not at all acceptable; 7 = very acceptable) for Constant and Varied Conditions.

Condition	Pilot Role	MR1-Pilot Verbal Delay Mean (SD)	MR1 ATCo Acceptability Mean Rating	MR2- Pilot Execution Delay Mean (SD)	MR2 ATCo Acceptability Mean Rating
Constant: VS- ES	UAS	2.07 (.50)	5.14	6.27 (2.52)	5.14
	Conventional	0.81 (.23)	5.00	3.24 (2.21)	5.00

Constant: VS-EL	UAS	2.12 (.27)	4.86	10.57 (1.32)	5.00
	Conventional	0.80 (.35)	5.43	4.79 (1.48)	5.43
Constant: VL-ES	UAS	5.52 (.32)	4.00	7.17 (1.43)	4.43
	Conventional	0.86 (.34)	4.43	5.56 (3.48)	5.29
Constant: VL-EL	UAS	5.43 (.25)	4.43	9.98 (2.35)	4.43
	Conventional	1.15 (.51)	4.43	4.97 (2.46)	5.29
Variable (averaged across the 4 variable scenarios)	UAS	3.72 (1.82)	5.07	8.98 (1.90)	5.18
	Conventional	0.94 (.54)	5.43	5.58 (2.49)	5.71

3.2 Tests of hypotheses and response to research questions

In what follows, we present the results of our simulation, organized around the five general questions outlined above.

(1) What is the effect of delays in verbal responding by UAS pilots on the ATCo acceptability ratings of the resulting MR1 components of all aircraft in the sector?

To analyze the effects of our manipulation of pilot verbal delays on the acceptability ratings of controllers, the ATCo acceptability ratings of Verbal latencies in the constant conditions were analyzed with a 3-factor repeated-measures ANOVA with the following factors: UAS-Verbal response delay (long vs. short), UAS-Execution delay (long vs. short) and pilot role (UAS vs. conventional aircraft). We found a significant main effect of UAS-Verbal communication delay, $F(1,12) = 7.49$; $p = .018$, $\eta_p^2 = .38$. Consistent with Hypothesis 1 above, we found that the mean acceptability rating for Short-Verbal delay conditions ($M = 5.1$ s; $SEM = 0.24$ s) was higher than for the Long-Verbal delay conditions ($M = 4.3$ s; $SEM = 0.32$ s). However, in contrast to Hypothesis 2, there was no significant main effect of pilot role, $F < 1.0$, which is surprising because short and long delays were inserted only in the UAS Verbal

communications and execution initiations. This finding indicates that longer latencies in UAS verbal responses affected ATCos' acceptability ratings of all communication latencies in a scenario, not just those of UAS pilots.

In *Figure 2*, the distributions of acceptability ratings are shown. These are separated by Long and Short verbal delays, collapsing across execution response times in the Constant Conditions and averaged across the four scenarios for the Variable Latency Condition. ATCos' ratings of the UAS verbal latencies are shown in the left panel of *Figure 2*, and ATCos' ratings of the conventional aircraft verbal latencies are shown in the right panel.

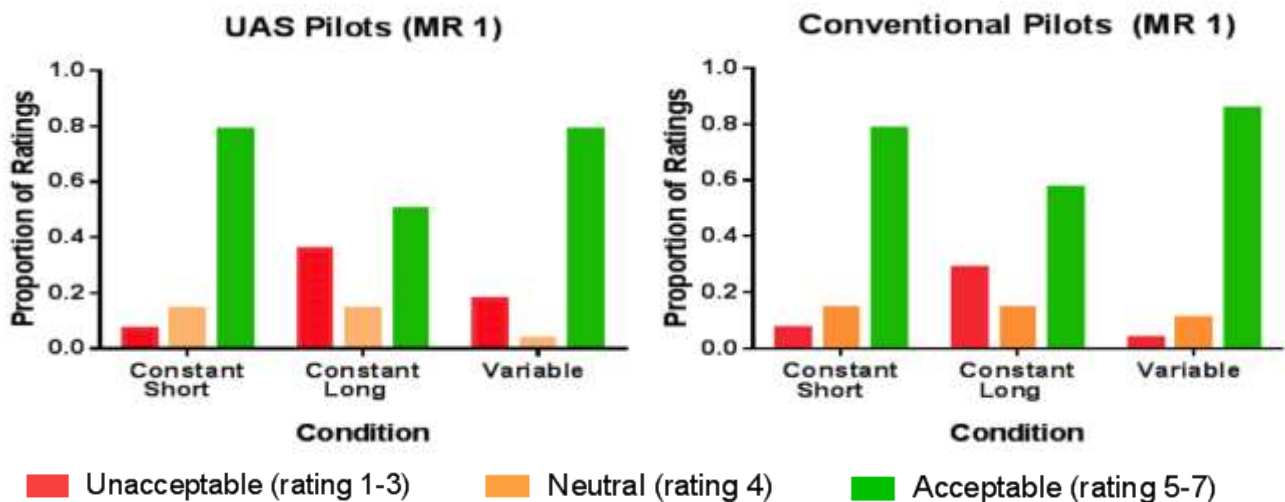


Figure 2. Distribution of ATCo Acceptability ratings (1-3 = unacceptable; 4 = neutral; 5-7 = acceptable) as a function of UAS Verbal Delay condition for MR1 of UAS (left panel) and conventional (right panel) pilots.

Overall, the majority of the ratings fell within an acceptable range of 5 or higher, showing that our controllers for the most part found these latencies manageable. For the Constant-Short Verbal Delay Condition, which led to an average latency of 2.10 s ($SEM = .09$ s) for UAS pilots and .80 s ($SEM = .09$ s) for conventional pilots, 78% of the ratings were in the acceptable range for UAS pilots (see *Figure 2* left panel) and 79% for conventional pilots (see *Figure 2* right panel). Thus, consistent with the ANOVA data, the frequency distribution of

ratings also indicates that the short delays were generally acceptable to ATCos. The fewest number of acceptable ratings occurred for Constant-Long verbal latencies, where the average latency was 5.48 s ($SEM = .09$ s) for UAS pilots and 1.01 s ($SEM = .09$ s) for conventional pilots. In this condition, only 49% of the ratings were acceptable for the UAS pilots and 57% for the conventional pilots. Thus, the long delays were only acceptable approximately half of the time.

(2) What is the effect of the predictability of the UAS-pilot verbal latencies (MR1) on ATCo acceptability ratings of MR1?

To evaluate the effect of predictability on ATCo acceptability ratings (Hypothesis 3), a two-factor repeated-measures ANOVA was run on verbal latency acceptability ratings, with the factors of Verbal-Delay Condition (Constant-Short, Constant-Long and Variable) and Pilot Role (UAS vs. conventional aircraft). A main effect of Delay Condition was obtained, $F(2,24) = 6.21$; $p = .007$, $\eta_p^2 = .34$. Post Hoc tests revealed that the acceptability rating was significantly lower for the Constant-Long ($M = 4.3$; $SEM = .32$) than Variable Delay Condition ($M = 5.3$; $SEM = .16$), $p = .042$. They also revealed a trend of higher acceptability for the Constant-Short ($M = 5.1$; $SEM = .24$) than Constant-Long Condition, $p = .054$. Thus, Hypothesis 3, which states that constant delays will be more acceptable than variable delays, was not supported. The finding that constant delays are not more acceptable than variable delays is also supported by the distribution of the acceptability ratings displayed in *Figure 2*. It shows that the acceptability ratings for the Variable-Delay Condition were similar to those of the Constant-Short Verbal Delay condition. For Variable-Delay scenarios, 78% of the ratings of UAS and conventional aircraft verbal response delays (which produced an average MR1 of 3.72 s and 0.92 s, respectively) were in the acceptable range. If predictability of verbal responses had negatively affected our controllers, we would have found the Variable-Delay Condition to produce the lowest proportion of

acceptable responses. Moreover, the fact that the percent of acceptable responses in the Variable-Delay Condition was comparable to the Constant-Short Verbal Delay Condition, and higher than for the Constant-Long Verbal Delay Condition suggests that the acceptability of verbal response latencies depended not only on the length of the latencies but also on the number of verbal responses having longer latencies: In the Variable-Delay scenarios, half of the verbal latencies were long, yet the distributions of acceptability ratings were consistent with the Constant-Short-Delay condition.

(3) What is the effect of delays in initiating command executions on the ATCo acceptability ratings of the resulting MR2 components of all aircraft in the sector?

To analyze the effects of our manipulation of UAS pilot execution initiation delays on the acceptability ratings of controllers, the ATCo acceptability ratings of Execution latencies for the constant delay conditions were analyzed with the 3-factor repeated-measures ANOVA described for MR1. In contrast to Hypothesis 4, which predicted higher acceptability ratings for the short execution delays, the analysis showed no significant main effect of Execution Delay Condition, $F < 1.0$. Likewise, contrary to Hypothesis 6, which predicts greater acceptability ratings for the conventional aircraft pilots than the UAS pilots, there was no significant effect of Pilot Role ($p > .31$). These factors also produced no significant interactions, $F_s < 1.22$, $p > .29$. Thus, our results regarding MR2 differ from those obtained with MR1. This does, however, support Hypothesis 5, which states that verbal delays are more salient to ATCos than delays in command execution. This makes sense given that controllers typically assume that the commands are executed promptly once clearances are acknowledged (Cushing, 1995).

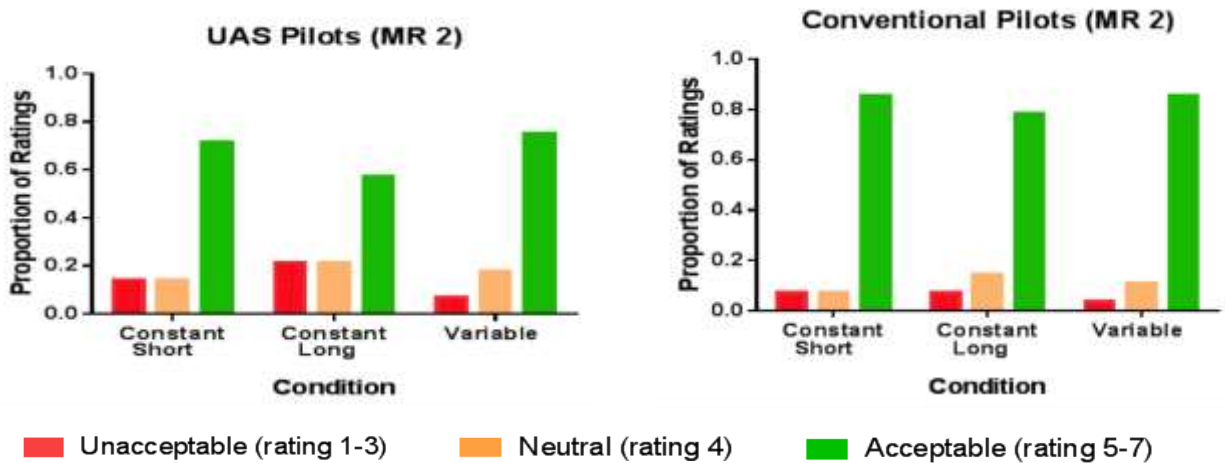


Figure 3. Distribution of ATCo Acceptability ratings (1-3 = unacceptable; 4 = neutral; 5-7 = acceptable) as a function of UAS Delay condition for MR2 of UAS (left panel) and conventional (right panel) pilots.

A descriptive analysis of the MR2 ratings allowed us to determine whether any of the delays in pilots' execution-initiation latencies were generally acceptable to ATCos. Similar to MR1, we examined the frequency distribution of ATCo acceptability ratings of UAS and conventional aircraft grouped into the categories of unacceptable, neutral, and acceptable, for MR2 (see *Figure 3*). As illustrated in the left panel of *Figure 3*, 71% of the UAS ratings were in the acceptable range for Constant-Short Execution Delays, which had a mean MR2 of 6.72 s ($SEM = .63$ s), and 56% of the ratings were in the acceptable range for Long-Execution Delays, which had a mean MR2 of 10.28 s ($SEM = .53$ s). Thus, although the differences in acceptability ratings were not as large as those observed for MR1, the short execution delays generally resulted in more acceptable ratings, though as we saw above, this was not statistically significant. For conventional aircraft (*Figure 3*, right panel), the distributions are quite similar. Acceptable responses were given for 92% of the scenarios in the Constant-Short Execution Delay condition, which had a mean MR2 of 4.4 s ($SEM = .63$ s). In the Constant-Long Execution Delay

condition, which had a mean MR2 of 4.8 s ($SEM = .54$ s), acceptable responses were given for 77% of the scenarios.

(4) What effect does the predictability of the UAS pilot execution initiation latencies (MR2) have on the ATCo acceptability ratings of MR2?

To examine the effect of predictability of the execution responses, a two-factor ANOVA similar to that conducted for MR1 was run on acceptability ratings of MR2. The main effect of pilot role was not significant, $F(1,12) = 1.66, p > .22$. Thus, hypothesis 6 was not supported. Moreover, in contrast to Hypothesis 7, which predicted that constant execution delays should yield greater acceptability ratings than variable ones, our ANOVA found no significant differences in acceptability between Constant and Variable scenarios ($p > .14$).

(5) What is the effect of verbal and execution response latencies on ATCo performance?

For each scenario and ATCo, we measured the number of LOS (a measure of sector safety), the average time of aircraft through the sector (a measure of efficiency), and the number of communication step-ons (which could affect both safety and efficiency). *Table 2* shows the means and standard errors (in parentheses) for each measure of sector performance for constant and variable scenarios.

Table 2. Mean (SEM) Sector Performance Measures for Constant and Variable Delay Scenarios

Performance Metric	Constant Delay	Variable Delay
Number of LOS	1.5 (.7)	2.1 (.7)
Distance Traveled (nm)	27.2(.54)	27.1(.42)
Number of Communication Step-Ons	47.4 (8.5)	47.5 (7.2)

To examine whether our manipulations affected sector performance of our ATCos in the constant scenarios, we first conducted a series of two-factor ANOVAs with the following

factors: UAS Verbal delay (short vs. long) and UAS Execution Initiation Delay (short vs. long). These analyses were conducted on the number of LOS, the mean distance travelled by aircraft, and the number of step-ons. There were no significant main effects or interactions for any of these variables (all $F_s < 1$). This goes against Hypothesis 8, which stated that performance should be better in the short delay conditions than in the long delay conditions. Moreover, in contrast to Hypothesis 9, which stated that performance should be better in the constant delay conditions than in the variable delay conditions, the predictability of the delays also did not seem to affect our ATCos. This was determined by conducting a separate set of repeated measures ANOVAs, comparing the three verbal delay conditions (Constant-Short, Constant-Long, and Variable-Delay). We found no differences between them in the number of LOS, in the time of aircraft through the sector, or in the number of step-ons (all $F_s < 1$). The same results were obtained in a set of repeated measures ANOVAs comparing our three execution initiation delay conditions, i.e., Constant-Short, Constant-Long, and Variable-Delay (all $F_s < 1$).

(6) What, if any, other factors influence ATCo ratings of the acceptability of delays in verbal responding and command execution?

It is possible that ATCos rate the acceptability of specific latencies based on more than the length of the delays themselves. To determine if ATCos used performance and sector outcomes to determine the acceptability of communication and execution latencies, Pearson correlations were computed between the acceptability ratings and measures of performance separately for constant and variable delay conditions. *Table 3* shows the relationships between acceptability ratings and performance for the constant delay conditions. Beginning with the ratings of the UAS response latencies, the correlation between Verbal Delay acceptability and MR1 showed a trend, such that shorter latencies tended to be associated with higher acceptability

ratings. The acceptability rating of Verbal Delay was also significantly correlated with the number of step-ons and average distance through the sector: as the number of step-ons and distance traveled decreased, the acceptability ratings increased. There was no significant correlation between Execution Initiation Delay acceptability and MR2. This is surprising because verbal delays and execution delays were crossed in our experiment. However, it may have been difficult for controllers to detect differences in the latency of command execution during the scenarios and their rating of the acceptability of execution delays may have been based instead on delays in verbal responding. Indeed, neither rating of acceptability was correlated with MR2.

Table 3. Pearson Correlations (*p* value in parentheses) Between Acceptability Ratings, Response Latencies and Sector Metrics for Constant Delay Conditions

<i>ATCo Ratings</i>	Verbal Resp. Latency (MR1)	Exec. Resp. Latency (MR2)	Number Step- Ons	Num. LOS	Ave. Dist. Travelled
<i>UAS</i>					
Acceptability Rating of MR1	-.35 (.07)	-.04 (.83)	-.46 (.01)	-.09 (.67)	-.37 (.05)
Acceptability Rating of MR2	-.34 (.08)	.20 (.32)	-.32 (.10)	-.007 (.97)	-.21 (.29)
<i>Conventional AC</i>					
Acceptability Rating of MR1	-.30 (.13)	-.02 (.93)	-.54 (.003)	-.40 (.04)	-.61 (.001)
Acceptability Rating of MR2	-.18 (.36)	-.03 (.87)	-.48 (.009)	-.27 (.17)	-.42 (.03)

With respect to the acceptability ratings of verbal delays and execution initiation delays of conventional-aircraft, neither was correlated with MR1 or MR2, the actual latencies in verbal responding and command execution observed in our experiment. However, the acceptability ratings of verbal response latencies were negatively correlated with the number of step-ons, number of LOS, and average distance aircraft traveled through the sector: as these metrics

decreased (performance increased), acceptability ratings increased. In terms of the acceptability rating of conventional aircraft execution delays, these were negatively correlated with the number of step-ons and with the average distance traveled by aircraft through the sector: as the number of step-ons and distance traveled decreased, the acceptability ratings increased. The fact that this was obtained with conventional aircraft and not with the UAS may reflect the fact that none of the LOS involved the UAS.

Table 4. Pearson Correlations (*p* value in parentheses) Between Acceptability Ratings and Response Latency for Variable Delay Conditions

<i>ATCo Ratings</i>	Verbal Response Latency (MR1)	Execution Response Latency (MR2)	Total Number Step Ons	Num LOS	Ave Dist Travelled
	<i>UAS</i>				
Acceptability Rating of MR1	-.18 (.36)	.21 (.28)	.04 (.86)	.10 (.61)	-.34 (.15)
Acceptability Rating of MR2	-.05 (.81)	.29 (.14)	-.29 (.14)	-.01 (.96)	-.12 (.56)
	<i>Conventional AC</i>				
Acceptability Rating of MR1	-.05 (.81)	-.29 (.13)	-.26 (.18)	.004 (.98)	-.08 (.67)
Acceptability Rating of MR2	-.40 (.04)	-.50 (.007)	-.43 (.02)	-.17 (.40)	-.18 (.35)

For the variable latency conditions, UAS acceptability ratings of verbal responses and execution initiation were not correlated with MR1, MR2 and all sector metrics as shown in *Table 4*. For conventional aircraft, the acceptability ratings of verbal delays were not related to either response component, but the acceptability ratings of execution initiation latencies were related to MR1 and MR2: as the MR 1 and MR 2 latencies decreased, acceptability ratings increased. Acceptability of execution initiation latencies were also negatively correlated with the number of step-ons.

IV. General Discussion

UAS integration into the NAS will require UAS-pilot responses to be similar to pilots of conventional aircraft (Askelson et al., 2013). This includes the latencies with which pilots verbally respond to ATCo commands, and the latencies with which they execute them. The present study inserted 1.5 s or 5 s delays to the verbal communications and command executions of UAS pilots and examined the effects of these manipulations on ATCo performance, and on their acceptability ratings of response latencies by UAS and conventional-aircraft pilots. In what follows, we discuss how our results pertain to each of our research questions and the corresponding hypotheses.

4.1 Effects of delays in verbal responding (MR1) by UAS pilots on ATCo acceptability ratings and air traffic management performance

Consistent with Hypothesis 1, we found that the short delays added to the UAS verbal response latencies, which led to an average UAS MR1 of 2.1 s, were more acceptable than the long added delays that led to an average UAS MR1 of 5.48 s. Furthermore, we found that the short MR1 latencies were generally acceptable, while the long ones were acceptable only in approximately half of the scenarios. Unfortunately, we are not aware of any operational data of UAS MR1 and MR2 components for comparison. As such, future studies need to capture the MR components of UAS in an operational environment.

Hypothesis 2 predicted that the acceptability ratings should be higher for conventional aircraft pilot verbal response latencies than for UAS pilots. This is because communication delays were only added to the responses of the latter. Indeed, the MR1 response times for the conventional pilots were much shorter than UAS pilots, being .81 s and 2.1 s in the short constant delay condition, respectively, and 1.01 s and 5.48 s in the constant long delay condition,

respectively. The response times of our conventional pilots were similar to those reported by Smith (2008) in a TRACON environment, but shorter than observed by Cardosi (1993) in an en route environment. It is not surprising that our MR1 latencies for conventional pilots are closer to the values obtained by Smith (2008), because ZLA-20 is a transition sector and our pseudopilots were not subjected to other real-world disruptions in their current task flow that could also contribute to longer verbal latencies. Nonetheless, the acceptability ratings for the two types of pilots did not differ. This result goes against Hypothesis 2. Finding no difference between the two types of pilots suggests that when ATCos judged acceptability of delays, their judgments reflected the broader operational context, a topic that will be explored further below.

Another way of viewing these results, however, is that the ATCos were more tolerant of delays in responding by UAS pilots than conventional aircraft pilots. If they had been equally stringent on both pilot types, the acceptability ratings for UAS pilots should have been much lower given that UAS response latencies were two to five times slower than conventional aircraft. One possible reason for this is that our ATCos may have assumed that the UAS pilots were less familiar with NAS operations, and may have expected them to take longer in responding, much like controllers often assume that new pilots will take longer to respond to their verbal commands, or will make more mistakes in doing so (Cushing, 1995). It is also possible that the ATCos assumed that the differences in response times reflect characteristics of the UAS control and communication interface that pilots cannot do anything about.

With respect to air traffic management, we found that the manipulation of UAS verbal delays did not affect the performance of ATCos in terms of LOS, mean distance traveled by aircraft, or the number of communication step-ons. Thus, in terms of our safety and efficiency metrics, our ATCos were able to adapt to the imposed verbal delays in responding by a single

UAS. This goes against Hypothesis 8, which predicted that performance would be better in the short verbal delay condition. Thus, our ATCos' performance remained consistent regardless of our manipulations, likely due to the expertise of our participants.

Another possible reason for why we failed to find differences in performance between the short and long pilot verbal delay conditions is because our simulation did not require multiple exchanges between the pilots and ATCos, unless step-ons occurred. Indeed, a study by Rantanen, McCarley, and Xu (2004) suggests this might be an important factor. In their study, they tested the effects of varied audio system latencies (i.e., set up and propagation times) and pilot latencies on ATCo performance. They examined these delays in part-task simulations that differed in terms of the number of communication exchanges that were required between pilots and ATCos to carry-out the actions. The tasks either required one communication exchange (i.e., ATCo issues clearance and pilot verbally acknowledges) or multiple communication exchanges (i.e., the task requires ATCos and pilots take more than one conversational turn). They found that both audio system and pilot latencies negatively affected ATCo performance only when multiple exchanges were required. In these conditions, greater audio-system latencies led to decreased lateral separation between aircraft and increased communication time. Higher pilot latencies were associated with lower lateral separation between aircraft, greater communication times, and a higher number of step-ons. In our study multiple exchanges were not required because our pseudopilots did not engage in negotiations with ATCos, but rather were required to be compliant with their commands. In sum, although it appears as though the expert controllers in our study were able to compensate when verbal delays were long, it remains to be seen whether this would continue to be the case if multiple communication exchanges were required, or if the sector workload was increased by having multiple UASs.

4.2 Effects of delays in command executions (MR2) by UAS pilots on ATCo acceptability ratings and air traffic management performance

Regarding latencies in execution initiation, Hypothesis 4 stated that acceptability ratings should be higher for the short-delay condition than the long condition. In contrast, we found that adding a short execution initiation delay to the UAS, which led to an MR2 of 6.72 s, did not yield significantly higher acceptability ratings than adding a long execution initiation delay, which led to an MR2 of 10.28 s. The results are consistent, however, with Hypothesis 5, which states that execution initiation delays should have less of an impact on ATCos than delays in verbal responding. In the debriefing session, ATCos indicated that they were not as concerned with the speed of the UAS execution because it was traveling at a slow speed of 110 knots, and changes in the UAS flight plan would not have a major impact on the traffic flows in a short period of time. This reason is consistent with the fact that, contrary to Hypothesis 9, there was no difference in ATCo performance between the short and long execution delay conditions. Also, because most of the conventional aircraft were at speeds of 250 knots or higher, the ATCos moved the faster moving traffic in the sector when necessary, and created a “buffer” zone around the UAS. The ATCos did note that they were able to create this buffer zone because there was only one UAS in the sector. They indicated that the execution delay would likely have played a bigger role on their performance if there were multiple UAS or if the UAS was much faster.

Furthermore, although the delays in executing commands were only added to the UAS-pilot responses, the acceptability ratings for the conventional aircraft pilots were similar to those for the UAS pilots. This finding goes against Hypothesis 6, which states that MR2 acceptability ratings should be higher for conventional pilots and is similar to what we observed with the acceptability ratings for pilot verbal response latencies. The comparable ratings for both UAS

and conventional pilots could have been a result of collecting the acceptability ratings after the scenario, which is a limitation of the present study, but one done out of necessity to accurately capture the MR components. Collecting the ratings after the scenario may allow ATCos to base their acceptability ratings of any given MR component on broader sector characteristics, and not just the length of the delay itself.

4.3 Effects of the predictability of delays in verbal responding and command execution and its effect on ATCo acceptability ratings and performance

In this study, we manipulated whether UAS verbal and execution delays were constant or variable within scenarios. With respect to the ATCo acceptability ratings, Hypotheses 3 and 7 stated that ATCo acceptability ratings should be higher in the constant-verbal-delay conditions than in the variable-verbal-delay conditions for MR1 and MR2, respectively. Moreover, with respect to performance, Hypothesis 9 predicted ATCo performance should be better in the constant delay conditions than in the variable delay conditions. The reason is that when delays are constant, operators can develop strategies for handling the delays. This may not be possible when the delays are unpredictable. In support of this reasoning, Watson, Walker, Ribarsky and Spaulding (1998) found that in a virtually controlled system when the delays had higher standard deviations, performance was degraded compared to when the standard deviation was smaller. Likewise, Lane et al. (2002) found that a system delay that was short but variable negatively affected operator performance more than a delay that was longer but constant.

However, in the present study, we did not find differences in acceptability or in ATCo performance between constant and variable delay conditions. Our controllers were able to compensate when verbal delays and execution delays were unpredictable. This ability is likely due to their many years of experience in air traffic management. It is an open question, of

course, whether unpredictable delays would be more disruptive to ATCos with less expertise. It is also important to investigate whether an effect of delay predictability would arise if multiple communication exchanges occurred between controllers and pilots. As noted above, Rantanen et al. (2004) found that pilot verbal delays were more disruptive of performance when multiple exchanges were required between ATCos and pilots. Under these conditions, it would be more important for controllers to be able to predict the length of the delays, especially the delays in verbal responding by UAS pilots, particularly if the delays vary within the same conversation between pilots and controllers. The predictability of the delays would also likely be more important if the controllers were managing multiple UAS in the sector. In this case, it would be more difficult for controllers to keep track of the delays that are at work in the responding of each UAS, as they could not assume that every UAS has the same response characteristics.

4.4 What factors are considered when ATCos rate acceptability of pilot verbal delays and execution initiation delays?

In the present study we showed that the overall acceptability rating depended on the number of longer vs. shorter latencies, with fewer longer latencies leading to higher ratings. This is because our variable latency conditions were rated equally acceptable as the scenarios that have constant short latencies, and the former have fewer of the long latencies than the constant-long scenarios. ATCos determined acceptability of a delay within a scenario from the standpoint of all aircraft, as we found very little difference in the ratings of UAS vs. conventional aircraft, though delays were added only to the UAS responses. Even though the ATCos provided the acceptability ratings separately for the UAS and conventional aircraft, it was done post-scenario, and this could have led the highly correlated ratings. Finally, the acceptability ratings were based on several performance factors in addition to the actual latencies produced by the UAS,

including the number of LOS, number of step-ons, and mean distance conventional aircraft travelled through the sector. Higher performance on these measures was related to higher acceptability ratings. Thus, when studying measured responses of UAS in the NAS it is therefore critical that a system-wide perspective be adopted, examining how ATCos interact with other aircraft as well (Vu & Chiappe, in press).

4.5 Implications

A key problem facing the integration of UAS in the NAS is specifying what is an acceptable measured response regarding both, pilot verbal responses (MR1) and initiation of commands (MR2). The present study begins to address this issue by comparing short and long delays in each of these two measured response components. An important implication is that in sectors with characteristics like the one examined in our study, delays in UAS pilot verbal responding that are approximately 1.5 s longer than those of conventional pilots are likely to be judged generally acceptable to ATCos. Delays in verbal responding that are approximately 5 s longer than those of conventional pilots are less likely to be acceptable to ATCos.

Of course, whether systems that allow for a delay that is up to 1.5 s longer than conventional can be implemented into the airspace is not a decision that can be answered solely by considering the subjective judgments of ATCos, or their performance in our simulation. The decision to do so will require simulations that take into account a more detailed analysis of the airspace under consideration, performance models of conventional aircraft, and UAS performance characteristics. Only in this way can one ensure that proper levels of safety and efficiency can be achieved. The same is the case with delays in executing clearances by UAS pilots. Although our study suggests that these are less relevant to ATCos than delays in verbal responding, this does not imply that other factors cannot be brought to bear in determining what

is acceptable in that regard. This issue therefore needs to be examined using more detailed simulations (e.g., Consiglio et al., 2008). What our study does tell us, however, is how ATCos are likely to react to the presence of UAS with these particular verbal and execution delays.

An ultimate goal of measured response research is to specify a threshold value beyond which performance falls below an acceptable level. By defining this threshold, requirements can be developed for systems and equipment so that the end-to-end measured response time does not exceed the threshold. Of particular relevance are systems and equipment involved with controlling a UAS and communicating with ATCos [communication speed and bandwidth, command and control data link speed and bandwidth, ground control station design (input time for manual vs. automated functions)]. Alternatively, or in combination, there may be limitations to systems and equipment that currently cannot be overcome, ones that lead to UAS measured response being unacceptable in certain phases of flight (e.g., approach and landing requiring more timely responses). If so, procedures and operational limitations may need to be implemented to safely integrate UAS. Our study is limited, however, in terms of its ability to specify an upper limit on what is an acceptable pilot verbal delay and execution delay. This is because it only implemented two possible response delays, 1.5 s and 5 s. Although we found that the former generally led to acceptable MR1 values while the latter did not, a range of values in between these two extremes would be required to identify a proper threshold. This will have to be examined in future studies.

To conclude, we found that delays in verbal responding are more salient to ATCos than delays in commencing to execute a command. Although ATCos were able to manage traffic safely and efficiently despite our manipulated delays, we did find that UAS pilot short verbal latencies of approximately 2 s are generally acceptable to ATCos. We also found that ATCo

ratings of the acceptability of delays are complex judgments that reflect multiple features of the airspace they are managing. Taking into account these characteristics will be essential in identifying an upper limit in what is an acceptable measured response.

References

- Agbolosu-Amison, S., Millner, D., Baden Jr., B., Coville, G., & Mondoloni, S. (2012). *Examining arrival time predictability in the NAS using monte carlo methods*. The MITRE Corporation Technical Report.
- Ambrosia, V., Cobleigh, B., Jennison, C., & Wegener, S. (2007). Recent experiences with operating UAS in the NAS. In: *AIAA Infotech Aerospace 2007 Conference and Exhibit*. Rohnert Park, California.
- Apaza, R., & Kubat, G. (2012). UAS in the NAS – control and non-payload communications modeling and simulation approach. *Proceedings of the 2012 Integrated Communications, Navigation and Surveillance Conference (ICNS 2012)*, May 2012.
- Askelson, M. A., Dreschsel, P., Nordlie, J., Theisen, C., Carlson, C., Woods, T., Forsyth, R., & Heitman, R. (2013). MQ-9 unmanned aircraft responsiveness to air traffic controller commanded maneuvers: Implications for integration into the national airspace system. *Air Traffic Control Quarterly*, 21, 79-92.
- Blickensderfer, B., Buker, T. J., Luxion, S. P., Lyall, B., Neville, K., & Williams, K. W. (2012). The design of the UAS ground control station: Challenges and solutions for ensuring safe flight in civilian skies. In *Proceedings of the Human Factors and Ergonomics Society 56th Annual Meeting* (pp.51-55). Santa Monica, CA: Human Factors and Ergonomics Society.
- Cardosi, K. M. (1993). Time Required for Transmission of Time-Critical Air Traffic Control Messages in an En Route Environment. *The International Journal of Aviation Psychology*, 3(4), 303–314.

- Clark, A. (2003). *Natural born cyborgs: Minds, technologies and the future of human intelligence*. New York: Oxford University Press.
- Consiglio, M., Hoadley, S., Wing, D., Baxley, B., & Allen, D. (2008). Impact of Pilot Delay and Non-Responsiveness on the Safety Performance of Airborne Separation. *Proceedings of the 8th AIAA ATIO Conference*, September 2008.
- Cushing, S. (1995). Pilot-air traffic control communications: it's not (only) what you say, it's how you say it. *Flight Safety Foundation*, 14, 25-31.
- Dillingham, G. L. (2013). *Unmanned aircraft systems: Continued coordination, operational data, and performance standards needed to guide research and development* (Report no. GAO-13-346T). Washington, DC: Government Accountability Office.
- Fern, L., & Shively, R. J. (2009). A comparison of varying levels of automation on the supervisory control of multiple UASs. In *Proceedings of AUVSI's Unmanned Systems North America 2009*, Washington, D.C..
- Ferrell, W. R. (1965). Remote manipulation with transmission delay. *IEEE Transactions on Human Factors in Electronics*, 6, 24–32.
- Hart, S. G., & Staveland, L. E. (1987). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 139–183). Amsterdam, The Netherlands: Elsevier.
- Held, R., Efstathiou, A., & Greene, M. (1972). Adaptation to displaced and delayed visual feedback from the hand. *Journal of Experimental Psychology*, 72, 887-891.
- Helleberg, J., & Wickens, C. D. (2003). Effects of Data-Link Modality and Display Redundancy on Pilot Performance: An Attentional Perspective. *The International Journal of Aviation Psychology*, 13, 189–210.

- Hopkin, V. D. (1995). *Human Factors in Air Traffic Control*. New York: Taylor & Francis.
- ICAO. (2011). *Unmanned Aircraft Systems (UAS)* (ICAO Cir 328). Montreal, Canada: International Civil Aviation Organization.
- Kamienski, J., Simons, E., Bell, S. & Estes, S. (2010). Study of unmanned aircraft systems procedures: Impact on air traffic control. *Digital Avionics Systems Conference*, IEEE/AIAA.
- Kerns, K. (1991). Data-Link communication between controllers and pilots: A review and synthesis of the simulation literature. *The International Journal of Aviation Psychology*, 1, 181-204.
- Lane, J. C., Carignan, C. R., Sullivan, B. R., Akin, D. L., Hunt, T., & Cohen, R. (2002). Effects of time delay on telerobotic control of neutral buoyancy vehicles. *Proceedings of IEEE International Conference on Robotics and Automation*, 3, 2874-2879.
- McCarley, J. S., & Wickens, C. D. (2005). *Human Factors Implications of UAVs in the National Airspace* (Technical Report AHFD-05-05/FAA-05-01). Savoy, IL: University of Illinois, Institute of Aviation, Aviation Human Factors Division.
- Merlin, P. W. (2009). *Ikhana unmanned aircraft system Western states fire missions*. Washington, DC: National Aeronautics and Space Administration.
- Merlin, P. W. (2013). *Crash course: Lessons learned from accidents involving remotely piloted and autonomous aircraft*. Washington, DC: National Aeronautics and Space Administration.

- Nadler, E., Mengert, P., DiSario, R., Grossberg, M., & Spanier, G. (1993). Effects of satellite- and voice-switching-equipment transmission delays on air traffic control communications. *The International Journal of Aviation Psychology*, 3, 315-325.
- Pierce, R. S. (2012). The effect of SPAM administration during a dynamic simulation. *Human Factors*, 54, 838-848.
- Prevot, T. (2002). Exploring the many perspectives of distributed air traffic management: The Multi Aircraft Control System: MACS. *International Conference on Human-Computer Interaction in Aeronautics, HCI-Aero 2002*. Cambridge, MA: MIT Press.
- Prevot, T., Homola, J.R., Martin, L.H., Mercer, J.S., & Cabrall, C.D. (2012). Toward automated air traffic control - investigating a fundamental paradigm shift in human/systems interaction. *International Journal of Human-Computer Interaction*, 28(2), 77-98.
- Rantanen, E. M., McCarley, J. S., & Xu, X. (2004) Time delays in air traffic control communication loop: Effect on controller performance and workload. *The International Journal of Aviation Psychology*, 14(4), 369-394.
- Sheridan, T. (1992). *Telerobotics, automation, and human supervisory control*. Cambridge, MA: MIT Press.
- Sheridan, T. (1993). Space teleoperation through time delay: Review and prognosis. *IEEE Transactions on Robotics and Automation*, 9, 592-606.
- Shively, R. J., Vu, K.-P. L., & Buker, T. J. (2013). Unmanned aircraft system response to air traffic control clearances: Measured response. In *Proceedings of the Human Factors and Ergonomics Society 57th Annual Meeting* (pp.31-35). Santa Monica, CA: Human Factors and Ergonomics Society.

- Smith, E. C. (2008). Analysis of controller-pilot communication performance in Area Navigation (RNAV) and conventional arrival operations. *Digital Avionics Systems Conference*, IEEE/AIAA.
- Sollenberger, R. L., McAnulty, D. M., & Kerns, K. (2003). *The effect of voice communications latency in high density, communications-intensive airspace* (Report no. DOT/FAA/CT-TN03/04). Atlantic City International Airport: Federal Aviation Administration William J. Hughes Technical Center.
- Tatar, D., Foster, G. & Bobrow, D. (1991). Design for Conversation: lessons from Cognoter. *International Journal of Man-Machine Studies*, 34, 185-209.
- Taylor, R. M. (1990). Situational Awareness Rating Technique (SART): The development for a tool for aircrew systems design (AGARD-CP-478). In *Situation awareness in aerospace operations* (pp. 3/1–3/17). Neuilly Sur Seine, France: NATO-AGARD.
- Tullis, T., & Albert, B. (2008). *Measuring the user experience*. Boston, MA: Morgan Kaufmann Publishers.
- Vu, K.-P. L., & Chiappe, D. (in press). Situation awareness in human systems integration. In D. Boehm-Davis, F. Durso, and J. Lee (Eds) *APA Handbook of Human Systems Integration*. Washington, DC: American Psychological Association.
- Vu, K.-P. L., Morales, G., Chiappe, D., Strybel, T. Z., Battiste, V. Shively, J., & Buker, T. J. (2013). Influence of UAS pilot communication and execution delay on controller’s acceptability ratings of UAS-ATC interactions. *Proceedings of the 32nd Digital Avionics Systems Conference* (pp. 6D4-1- 4.D.4-13). Syracuse, NY: IEEE.
- Vu, K.-P. L., Strybel, T. Z., Battiste, V., Lachter, J., Dao, A.-Q. V., Brandt, S., Ligda, S., & Johnson, W. W. (2012). Pilot performance in trajectory-based operations under concepts

- of operation that vary separation responsibility across pilots, air traffic controllers, and automation. *International Journal of Human-Computer Interaction*, 28,107-118.
- Watson, B., Walker, N., Ribarsky, W., & Spaulding, V. (1998). Effects of variation in system responsiveness on user performance in virtual environments. *Human Factors*, 40, 403-414.
- Wickens, C. D., & Hollands, J. G. (2000). *Engineering psychology and human performance* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.
- Williams, K. (2007). *An assessment of pilot control interfaces for unmanned aircraft*. (Report no. DOT/FAA/AM-07/8). Washington, DC: Federal Aviation Administration.
- Williams, K. W. (2012). *An Investigation of sensory information, levels of automation, and piloting experience on unmanned aircraft pilot performance* (Report no. DOT/FAA/AM-12/4). Washington, DC: Federal Aviation Administration.
- Wourms, D. F., Ogden, J., & Metzler, T. R. (2001a). *NEXCOM auditory latency, Volume I: Noncopyrighted literature search results* (Report No. HSIAC–SS–2001–005). Wright Patterson Air Force Base, OH: The Human Systems Information Analysis Center.
- Yang, J., Rantanen, E. M., & Zhang, K. (2010). The impact of time efficacy on air traffic controller situation awareness and mental workload. *International Journal of Aviation Psychology*, 20, 74-91.
- Ziccardi, J., Roberts, Z., O'Connor, R., Rorie, C., Morales, G., Battiste, V., Strybel, T. Z., Chiappe, D., Vu, K.-P. L., & Shively, J. (2013). Measuring UAS pilot responses to common air traffic clearances. In S. Yamamoto (Ed.): HIMI/HCI 2013, Part II, *Lecture Notes in Computer Science*, 8017, 606-612.

Authors Note

Acknowledgements

This project was supported by NASA cooperative agreement NNX12AH23A, NASA UAS in the NAS (Jay Shively and Walter Johnson, Technical Monitors), and NASA cooperative agreement NNX09AU66A, “Group 5 University Research Center: Center for Human Factors in Advanced Aeronautics Technologies” (Brenda Collins, technical monitor).

Authors’ Brief Bio

Kim-Phuong Vu received her PhD in Cognitive Psychology at Purdue University in 2003. She is a Professor in Human Factors Psychology at California State University, Long Beach and the Associate Director of the Center for Human Factors in Advanced Aeronautics Technologies, a NASA Group 5 University Research Center.

Greg Morales is a graduate student in the Human Factors program at California State University Long Beach. He anticipates on receiving his MS degree in Human Factors Psychology in Summer 2014.

Dan Chiappe received his PhD in Cognitive Psychology at the University of Toronto in 1997. He is a Professor in Human Factors at California State University, Long Beach and a Research Associate of the Center for Human Factors in Advanced Aeronautics Technologies.

Thomas Strybel received his PhD in Psychology at Arizona State University in 1986. He is a Professor in Human Factors Psychology at California State University, Long Beach and the Director of the Center for Human Factors in Advanced Aeronautics Technologies.

Vernol Battiste received his MA degree in Psychology at San Jose State University in 1987. He is principal investigator and co-lead of NASA Ames’s Flight Deck Display Research Laboratory, a group investigating advanced cockpit displays of traffic information.

Robert Jay Shively received his MS degree in Cognitive Psychology at Purdue University in 1984. He is the Lead Project Engineer for the Human System Integration Subproject for the NASA UAS Integration into the NAS project and a Co-Lead of NASA Ames's Flight Deck Display Research Laboratory.

Timothy J. Buker received his MS in aviation human factors and BS in aerospace engineering, both from Florida Institute of Technology, in 2009 and 2007, respectively. At the time of this research, he was a human factors specialist with SAIC, working at the Federal Aviation Administration Unmanned Aircraft Systems Integration Office in Washington, D.C.